

# Correct, Robust, and Useful: I can haz all three?



**MARCIN PORWIT**  
**CSE577 FINAL PROJECT**  
**FALL 2011**

# Aims



- **The What:**

- Globally identify any image
- Uniquely identify each user
- Embed information directly in image



- **The Why:**

- To know the origin of an image
- To know the identity of the researcher
- **To make this automatic and painless -> pervasive**



# The Nitty Gritty



- **User Identity**
  - X.509 Certificate
    - ✦ /DC=org/DC=cilogon/C=US/O=University of Washington/CN=MARCIN PORWIT A807
    - ✦ Trust delegation based on NSF CILogon Service (<https://cilogon.org>)
- **Image Identifier**
  - Place name-based UUID, according to RFC4122
    - ✦ 21015C3-EBCC-4C60-BE8B-A9CB63BED091
    - ✦ Place == repository/host/data bank
    - ✦ UUID stored in DB, tied to image

# More Nitty Gritty



- Images

- ~~DICOM~~ NIFTI Brain MRI files

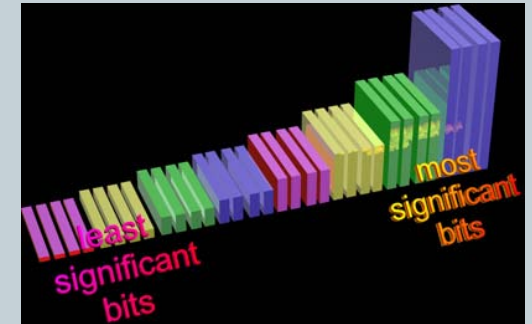
- ✦ 16-bit luminescence values, very limited headers
- ✦ neuroimaging standard



- Encoding

- ~~Haar wavelet~~ LSB steganography

- ✦ Stega-what?
- ✦ Figuring out what data is not important and hijacking that



# But does it deliver?



- **Correctness**

- + Plain ASCII
- + Shorter Messages
- - Error-correcting codes
- - Long Messages

- **Robustness**

- + OK for base NIFTI
  - ✦ > 50% recovery rate
- - Breaks down for subsequent ops
  - ✦ Processing for computer vision is destructive

- **Utility**

- + Does not affect computed result
- - Does not preserve payload

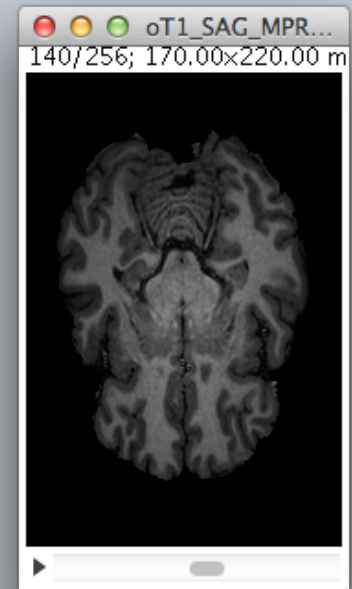
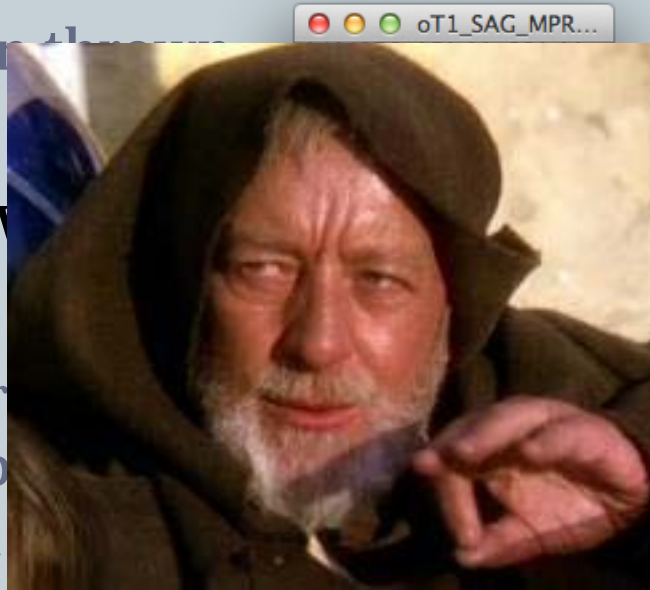
- **Preliminary results**

- Nolan is running a larger set of images and operations to confirm

# What now?



- **Obvious in retrospect**
  - Lot of information thrown away
- **Back to the drawing board**
  - Start with fully processed image and work backwards to figure out total



Not the results I was looking for